# Leveraging One-hop Information in Massive MIMO Full-Duplex Wireless Systems

Wenzhuo Ouyang, Jingwen Bai, Ashutosh Sabharwal, *Fellow, IEEE*

*Abstract*—We consider a single-cell massive MIMO full-duplex wireless communication system, where the base-station (BS) is equipped with a large number of antennas. We consider the setup where the single-antenna mobile users operate in half-duplex, while each antenna at the BS is capable of full-duplex transmissions, i.e., it can transmit and receive simultaneously using the same frequency spectrum. The fundamental challenge in this system is *intra-cell* inter-node interference, generated by the transmissions of uplink users to the receptions at the downlink users. The key operational challenge is estimating and aggregating inter-mobile channel estimates, which can potentially overwhelm any gains from full-duplex operation.

In this work, we propose a scalable and distributed scheme to optimally manage the inter-node interference by utilizing a *"one-hop information architecture"*. In this architecture, the BS only needs to know the signal-to-interference-plus-noise ratio (SINR) from the downlink users. Each uplink user needs its own SINR, along with a weighted signal-plus-noise metric from its one-hop neighboring downlink users, which are the downlink users that it interferes with. The proposed one-hop information architecture does not require any network devices to comprehensively gather the vast inter-node interference channel knowledge, and hence significantly reduces the overhead. Based on the one-hop information architecture, we design a distributed power control algorithm and implement such architecture using overheard feedback information. We show that, in typical asymptotic regimes with many users and antennas, the proposed distributed power control scheme improves the overall network utility and reduces the transmission power of the uplink users.

## I. Introduction

With each generation of wireless standards, the number of antennas at the infrastructure nodes has continued to grow to meet the increasing demand of mobile data. For example, both cellular and WiFi standards now support up to 8 antennas at infrastructure nodes. A promising and now standardized approach to use multiple antennas at the infrastructure is to use Multi-User Multiple Input Multiple Output (MU-MIMO) for supporting multiple uplink or downlink data streams in same time-slot. More recently, massive MIMO regime has been explored, which allows for a large number of antennas to reside at the infrastructure (with orders of magnitude more antennas compared to conventional MIMO systems), has attracted large interests in both academia and industry. The theoretical ([13], [18], [19]) and experimental results ([25], [27]) have shown that massive MIMO can lead to a number of desirable properties from the network design point of view, that include reduced inter-beam interference, improved energy efficiency, and reduced inter-cell interference, among others. As a result, massive MIMO is being considered for standardization as one of the key technologies in next-generation of

wireless systems in both cellular and WiFi (e.g., 3GPP LTE Release 12 [2] and 802.11-ax [3]). Channel models with 64 BS antennas have already been standardized [2], with research platform supporting even larger configurations; e.g. 96 BS antennas [26].

In addition to MU-MIMO, another potential avenue to achieve higher spectral efficiency is to leverage full-duplex transmission, where a full-duplex-capable device can transmit and receive at the same time using the same frequency spectrum. The full-duplex mode of network operation has the potential to double the spectrum efficiency of wireless networks and can bring substantial flexibility to higher layer design [9]. In fact, in-band full-duplex has already become part of the ongoing standard both in 3GPP [1] and 802.11-ax [3]. While it can be hard to integrate the full-duplex capabilities to client devices due to the processing and energy constraint, it is in fact feasible to design near-perfect full-duplex base-stations thanks to available freedom (bigger size, non-battery-powered operation) in their designs (e.g., see [11], [12], [22] and the references therein). One method to leverage full-duplex infrastructure with half-duplex mobile handsets is to have *simultaneous* multi-user up- and downlink transmissions. However, the potential for simultaneous up- and downlink MU-MIMO at the BS leads to a new challenge – the inter-node interference *within* each cell, i.e intra-cell interference, as the transmissions of uplink interfere with the receptions of downlink as illustrated in Figure 1. The inter-node interference hence poses a fundamental challenge to enabling full-duplex at the BS and needs to be managed efficiently.

In the presence of intra-cell interference, a centralized scheme can be used where the infrastructure device aggregates comprehensive knowledge of all the uplink, downlink and inter-node channel to make jointly optimal decisions on resource allocation. The centralized scheme may be useful for sparse scenario with small number of users. It will, however, incur a significant amount of overhead as the number of users grows, with the inter-node interference being *the dominant bottleneck*. For example, consider the setup with $M$ full-duplex-capable antennas[1] at the BS, and with $K_{ul}$ half-duplex single antenna uplink and $K_{dl}$ half-duplex single antenna downlink users, henceforth denoted as $(M, K_{ul}, K_{dl})$ massive-MIMO full-duplex system, in Figure 1. The centralized control scheme requires the full-duplex BS to aggregate knowledge of $K_{ul} \times K_{dl}$ inter-node interference channels, in addition to $M \times K_{ul}$ uplink and $M \times K_{dl}$ downlink channels. Note that there has been recent development on uplink/downlink channel estimation in Massive MIMO with reasonable complexity [17],

[1]A full duplex capable antenna is capable of transmitting and receiving at the same time. It can be realized by either using one antenna (e.g., [4]) or using a pair of antennas (e.g., [11][12]).

[31]. For example, the $M \times K_{ul}$ uplink channels can be estimated by letting $K_{ul}$ uplink users send pilots to the BS, and downlink channels can be obtained by letting $K_{dl}$ downlink users send pilots and estimated by utilizing channel reciprocity. Hence the $M \times (K_{ul} + K_{dl})$ up- and downlink channels can obtained with $(K_{ul} + K_{dl})$ transmissions of overhead (pilots). However, to obtain $K_{ul} \times K_{dl}$ inter-node interference channels, each of the $K_{ul} \times K_{dl}$ channel information needs to estimate at the uplink/downlink users and sent to the infrastructure; inter-node interference channels are two-hop away from BS. Hence acquiring all the channel knowledge of $K_{ul} \times$ inter-node interference channels incurs overhead at least the order of $K_{ul} \times K_{dl}$. As an example, under the scenario with $K_{ul} = 15$ uplink and $K_{dl} = 20$ downlink users, there can be as many as $K_{ul} \times K_{dl} = 300$ inter-node interference channels, which is significantly more than $K_{ul} + K_{dl} = 35$. Therefore, for centralized scheme, obtaining the inter-node interference becomes the main bottleneck and will render the overall centralized architecture non-scalable.

In this paper, we thus seek distributed methods for resource allocation that significantly reduce *who* needs to know *what* about *which* channels. We propose, to the best of our knowledge, the first scalable and distributed interference management techniques in massive MIMO full-duplex systems with half-duplex clients. Our contributions are listed as follows.

- We create a scalable information architecture, namely the *one-hop information architecture*, that specifies the network information needed in each cellular devices to perform optimal interference management. Specifically, we consider an architecture where (i) BS knows only SINR of downlink channels, much like current cellular systems which rely on periodically collected SINR information for power and rate control, and (ii) uplink transmitters know only about interference channel strength, along with a weighted interference-plus-noise (IN) metric, from the nearby downlink receivers. Therefore, for the $(M, K_{ul}, K_{dl})$ massive-MIMO full-duplex system, compared to a fully centralized decision, the proposed system requires only local, at most $K_{dl}$ one-hop information at each uplink user from its neighboring downlink receivers, thus avoiding the significant non-scalable overhead for aggregating the $K_{ul} \times K_{dl}$ interference channel information at the BS.
- Based on the one-hop information architecture, we propose a distributed power allocation algorithm to achieve the optimal sum-utility in the system. By analyzing the structural properties of the $(M, K_{ul}, K_{dl})$ massive-MIMO full-duplex system, we identify that the up- and downlink power control, which is coupled by inter-node interference, can in fact be carefully decoupled. Such a decoupled structure turns out to be an enabler for the one-hop information architecture and the optimal distributed algorithm.
- We further show that, in typical asymptotic regimes where the amount of antennas scales up with the number of users, distributed power allocation can not only improve the overall system utility, it can also asymptotically
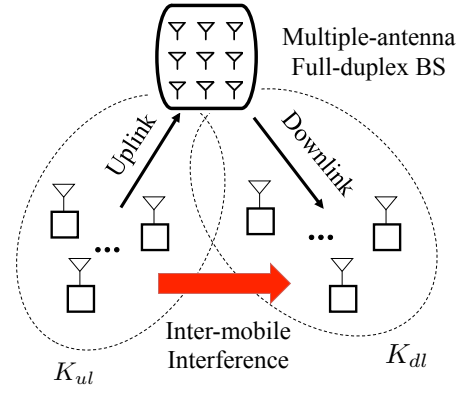


Fig. 1. Full-duplex multi-user network with $K_{ul}$ uplink users transmitting to the BS and $K_{dl}$ downlink users receiving from BS.

reduce the overall amount of transmission power and hence improve the energy efficiency of mobile users.

Note that, there can be scenarios with dense users and hence strong inter-node interference, where half duplex mode can be more efficient solution. The scope of this paper is not to specify when and where to use full duplex mode of operation, but rather to explore the optimal and scalable power control schemes for massive MIMO when the full duplex mode is in use.

The rest of the paper is organized as follows. In Section III we formulate the physical layer model of massive MIMO full-duplex system. We proceed to present the power control problem in Section IV. In Section V, we present the one-hop information architecture that facilitates us to propose distributed interference management algorithm. An overhearing-based scheme is presented in Section VI to implement the information architecture. We present asymptotic analysis in large number of antenna regime in Section VII. Numerical evaluation are provided in Section VIII followed by conclusion remarks in Section IX.

## II. RELATED WORKS

There has been extensive research on resource allocation in various interference-constrained wireless networks. The interference in the wireless networks is often captured by the SINR, which in turn determines the achievable transmission rate. Power control is an important mechanism to mitigate network interference and have been well studied (see survey [7]). Such problem often requires solving nonconvex optimization problem that is proven to be NP hard to solve [16].

One line of research has focused on centralized control where a centralized controller gathers all the information across the network to perform resource allocation and interference management. For instance, the authors in [15] proposed centralized power control algorithm under high SINR regime using geographic programming. Signomial programming based approaches are proposed in [8] to achieve a local optimal solutions. An iterative optimal algorithm was proposed in [20] to achieve global optimality for a subset of non-convex objective functions. A branch and bound algorithm was proposed in [29] to solve weighted sum-rate maximization

problems. Approximation algorithm was proposed in [28] to solve the NP-hard power control problem.

Another line of research has studied distributed power control algorithms for interference management. In the high SINR regime, distributed power control algorithm was converted to geometric programming [6]. Distributed price-based power control algorithm was proposed in [14] and shown to be optimal in certain scenarios. Distributed Power Control and scheduling based on back-pressure algorithm is proposed in [30] in high SINR regime. Recently, optimal distributed scheduling based on SINR model is proposed in to achieve transmission rate over a *power control region* [5].

Some recent work studied resource allocation for full duplex systems. The authors in [10] studied optimal subcarrier allocation for Full Duplex OFDMA network. A cell partitioning method was proposed in [24] that allocates frequency resources according to the partitioning of a single cell.

## III. SYSTEM MODEL

In this section, we describe the channel model for a single-cell massive MIMO full-duplex system where the BS is full-duplex capable while the mobile users operates in single antenna half-duplex mode. We assume the full-duplex BS is equipped with $M$ antennas to serve *both* uplink and downlink users in the same time/frequency slot. The sets of uplink and downlink users are denoted by $\mathcal{K}_{ul}$ and $\mathcal{K}_{dl}$ respectively, with $|\mathcal{K}_{ul}| = K_{ul}$ and $|\mathcal{K}_{dl}| = K_{dl}$, as shown in Figure 1.

For each uplink/downlink user $k$, an associated utility function $U_k(r)$ represents the utility it achieves under transmission rate $r$. It reflects the level of satisfaction achieved for each of the users at a given rate. Each utility function is assumed to be a non-decreasing, continuously differentiable, strictly concave function, where $U_k'(r)$ is the derivative of the utility function. The concavity property captures the diminishing return property of the utility achieved by the users in the network [23]. These utility functions are also used to capture different fairness requirement over the network. We are interested in controlling the transmission power, over both uplink and downlink in the massive MIMO full-duplex system, so that the achievable transmission rate provides the maximum overall sum-utility over the network. To characterize the physical-layer relationship between the transmission power and the corresponding achievable transmission rate, we start by presenting the channel model of the massive MIMO full-duplex system.

The propagation channel model in our system consists of two parts: small-scale fading due to multipath and large-scale fading due to path loss. The channels are represented by the set $\mathcal{G} = \{\mathbf{G}^{ul}, \mathbf{G}^{dl}, \mathbf{G}^I\}$, where each element in $\mathcal{G}$ denotes the propagation matrix of complex-valued channel coefficients for uplink, downlink and inter-node interference channels, respectively. We can express $\mathbf{G}^k$, $k \in \{ul, dl, I\}$ as

$$\mathbf{G}^k = \mathbf{L}^k (\mathbf{D}_{\mathbf{g}}^k)^{\frac{1}{2}}$$

where $\mathbf{L}^k$ represents the small-scale fading matrix with a magnitude of one for each entry, and $(\mathbf{D}_{\mathbf{g}}^k)^{\frac{1}{2}} = diag(\sqrt{\mathbf{g}^k})$ is a diagonal matrix whose diagonal entries constitutes a vector $\mathbf{g}^k$ of path loss coefficients.

### A. Uplink

At the BS with $M$ antennas, the received uplink signal $\mathbf{y}^{ul}$ at the BS is an $M \times 1$ vector given as

$$\mathbf{y}^{ul} = \mathbf{G}^{ul} \mathbf{x}^{ul} + \mathbf{z}^{ul}, \tag{1}$$

where $\mathbf{G}^{ul}$ is an $M \times K_{ul}$ uplink channel matrix and $\mathbf{x}^{ul}$ is a $K_{ul} \times 1$ vector, with each element representing the transmit signal of the uplink users. The vector $\mathbf{z}^{ul}$ represents the receiver additive noise, each entry is independent and identically distributed (i.i.d.) and drawn from a circularly symmetric complex Gaussian distribution with zero mean and variance of $N_0$ denoted as $\mathcal{CN}(0, N_0)$.

In the uplink, we let $\mathbf{P}^{ul} = [P_1^{ul}, P_2^{ul}, \cdots, P_{K_{ul}}^{ul}]$ denote the vector of transmit powers, where the $P_i^{ul}$ represents the transmit power of uplink user $i$. Each user is subject to a maximum instantaneous power constraint $P_i^{max}$, i.e., $P_i^{ul} \le P_{max}^{ul}$, $\forall i \in \mathcal{K}_{ul}$.

Following the techniques in [18], [21], the BS can simply process the received signal using maximum-ratio combining (MRC),

$$(\mathbf{G}^{ul})^H \mathbf{y}^{ul} = (\mathbf{G}^{ul})^H \mathbf{G}^{ul} \mathbf{x}^{ul} + (\mathbf{G}^{ul})^H \mathbf{z}^{ul}, \tag{2}$$

where superscript $H$ denotes conjugate transpose. We hence have

$$\frac{(\mathbf{G}^{ul})^H \mathbf{G}^{ul}}{M} = (\mathbf{D}_{\mathbf{g}}^{ul})^{\frac{1}{2}} \left( \frac{(\mathbf{L}^{ul})^H \mathbf{L}^{ul}}{M} \right) (\mathbf{D}_{\mathbf{g}}^{ul})^{\frac{1}{2}}, \tag{3}$$

where $(\mathbf{D}_{\mathbf{g}}^{ul})^{\frac{1}{2}}$ is a $K_{ul} \times K_{ul}$ matrix and $\mathbf{L}^{ul}$ is an $M \times K_{ul}$ matrix. In massive MIMO, the number of antennas at BS is much larger than that at mobiles. Thus when $M \gg K_{ul}$, under favorable propagation conditions as manifested in [18], the columns of the small-scale fading matrix are asymptotically orthogonal, the following convergence is obtained [18],

$$\left( \frac{(\mathbf{G}^{ul})^H \mathbf{G}^{ul}}{M} \right)_{M \gg K_{ul}} \to \mathbf{D}_{\mathbf{g}}^{ul} \tag{4}$$

Now substituting (4) into (2), we have

$$\left( \frac{(\mathbf{G}^{ul})^H \mathbf{y}^{ul}}{M} \right)_{M \gg K_{ul}} \to \mathbf{D}_{\mathbf{g}}^{ul} \mathbf{x}^{ul} + \frac{(\mathbf{G}^{ul})^H \mathbf{z}^{ul}}{M}. \tag{5}$$

We then compute the Signal-to-Interferene-and-Noise ratio (SINR) for uplink user $i$ as

$$\mathsf{SINR}_i^{ul} = \frac{M P_i^{ul} g_i^{ul}}{N_0}, \tag{6}$$

where $\sqrt{g_i^{ul}}$ is the $i$th entry in the diagonal matrix $(\mathbf{D}_{\mathbf{g}}^{ul})^{\frac{1}{2}}$. Note that in the above expression, we have used the asymptotic result in (5). Under the domain of massive MIMO, we assume the SINR expression is valid for large value of $M$, hence ignoring the negligible error term.

Therefore, with receiver MRC, the uplink user $i$ achieves a transmission rate

$$R_i^{ul}(P_i^{ul}) = \log \left( 1 + \frac{M \cdot P_i^{ul} g_i^{ul}}{N_0} \right), i \in \mathcal{K}_{ul}, \tag{7}$$

where, recall that, $g_i^{ul}$ is the corresponding path loss coefficient for $i$-th user.

## B. Downlink

The received signal at the $K_{dl}$ downlink users $\mathbf{y}^{dl}$ is a combination of the downlink signals and the interfering uplink signals, and is given by

$$\mathbf{y}^{dl} = \mathbf{G}^{dl}\mathbf{x}^{dl} + \mathbf{G}^I\mathbf{x}^{ul} + \mathbf{z}^{dl}, \tag{8}$$

where $\mathbf{y}^{dl}$ is a $K_{dl} \times 1$ received signal vector, $\mathbf{G}^{dl}$ is a $K_{dl} \times M$ downlink channel matrix and $\mathbf{x}^{dl}$ is an $M \times 1$ downlink transmit signal vector; $\mathbf{G}^I$ is a $K_{dl} \times K_{ul}$ inter-node interference channel matrix; $\mathbf{z}^{dl}$ is the receiver additive Gaussian noise which contains i.i.d. $\mathcal{CN}(0, N_0)$ entries.

In the downlink, we let $\mathbf{P}^{dl} = [P_1^{dl}, P_2^{dl}, \cdots, P_{K_{dl}}^{dl}]$ denote the vector where the element $P_j^{dl}$ represents the transmit power to downlink user $j$. The downlink transmission power allocation is subject to a constraint on the total amount of transmission powers, i.e., $\sum_{j=1}^N P_j^{dl} = P_{tot}^{dl}$.

Note that we assume perfect serf-interference cancellation at the full-duplex-capable BS. This is inline with recent results that self-interference cancellation is capable of suppressing self-interference below the noise floor [12].

Similar to the uplink, in massive MIMO downlink, under the conditions of $M \gg K_{dl}$, we have

$$\left( \frac{\mathbf{G}^{dl}(\mathbf{G}^{dl})^H}{M} \right)_{M \gg K_{dl}} \to \mathbf{D}_{\mathbf{g}}^{dl}. \tag{9}$$

Following the techniques in [21], BS can perform conjugate beamfoming precoding,

$$\mathbf{x}^{dl} = \frac{1}{\sqrt{M}}(\mathbf{G}^{dl})^H(\mathbf{D}_{\mathbf{g}}^{dl})^{-\frac{1}{2}}\mathbf{D}_p^{\frac{1}{2}}\mathbf{s}^{dl}. \tag{10}$$

where $\mathbf{D}_p^{\frac{1}{2}} = diag(\sqrt{\mathbf{P}^{dl}})$ is a $K_{dl} \times K_{dl}$ diagonal matrix. The $K_{dl} \times 1$ vector $\mathbf{s}^{dl}$ contains downlink messages where each message has unit power, i.e., $E(|s_j^{dl}|) = 1$ for all $j \in \mathcal{K}_{dl}$.

We substitute (9) and (10) into (8) and we obtain that

$$\left( \frac{\mathbf{y}^{dl}}{\sqrt{M}} \right)_{M \gg K_{dl}} \to (\mathbf{D}_{\mathbf{g}}^{dl})^{\frac{1}{2}}\mathbf{D}_p^{\frac{1}{2}}\mathbf{s}^{dl} + \frac{\mathbf{G}^I\mathbf{x}^{ul}}{\sqrt{M}} + \frac{\mathbf{z}^{dl}}{\sqrt{M}}. \tag{11}$$

Because the uplink and downlink transmissions are over the same frequency band, the uplink transmissions causes interference to the downlink receptions, as shown in Figure 1. The interference channel strength between uplink $i$ and downlink $j$ is denoted as $g_{ij}^I$. For the uplink user $i$, we let $\mathcal{N}_i \subset \mathcal{K}^{dl}$ denote the set of downlink users that receives non-negligible interference from user $i$. For downlink user $j$, we let $\mathcal{N}_j \subset \mathcal{K}^{ul}$ denote the set of uplink users that imposes non-negligible interference to user $j$.

In the massive MIMO scenario where $M \gg K_{dl}$, the SINR of downlink user $j$ can then be calculated as

$$\mathsf{SINR}_j^{dl} = \frac{M \cdot P_j^{dl} g_j^{dl}}{\sum_{i \in \mathcal{N}_j} g_{ij}^I P_i^{ul} + N_0} \tag{12}$$

where $\sqrt{g_j^{dl}}$ is the $j$th entry in the diagonal matrix $(\mathbf{D}_{\mathbf{g}}^{dl})^{\frac{1}{2}}$, and, similar to the uplink, we have ignored the negligible error term in (12).

We consider the setup where the downlink receivers treat the inter-node interference as noise. Therefore, with transmit conjugate beamforming, the downlink transmission rate to user $j \in \mathcal{K}_{dl}$ is expressed as

$$R_j^{dl}(P_j^{dl}, \mathbf{P}^{ul}) = \log\left( 1 + \frac{M \cdot P_j^{dl} g_j^{dl}}{\sum_{i \in \mathcal{N}_j} g_{ij}^I P_i^{ul} + N_0} \right), \tag{13}$$

which depends on both uplink and downlink power allocation schemes $\mathbf{P}^{ul}, \mathbf{P}^{dl}$. We henceforth denote the Interference-plus-Noise (IN) by $\mathsf{IN}_j = \sum_{i \in \mathcal{N}_j} g_{ij}^I P_i^{ul} + N_0$.

## IV. POWER CONTROL PROBLEM

In this section, we formally define the power control problem. The objective of power allocation is to control the transmission powers $\mathbf{P}^{ul}, \mathbf{P}^{dl}$ to jointly manage the inter-node interference and intelligently allocate the downlink transmission powers under the interference, so that the overall sum-utility in the system is maximized. Recall that $R_i^{ul}(P_i^{ul})$ and $R_j^{dl}(P_j^{dl}, \mathbf{P}^{ul})$ are defined in (7) and (13) respectively. The power allocation problem is defined as the following optimization problem.

$$\max_{\mathbf{P}^{ul}, \mathbf{P}^{dl}} \sum_{i \in \mathcal{K}_{ul}} U_i\left( R_i^{ul}(P_i^{ul}) \right) + \sum_{j \in \mathcal{K}_{dl}} U_j\left( R_j^{dl}(P_j^{dl}, \mathbf{P}^{ul}) \right) \tag{14}$$

$$s.t., \quad \sum_{j \in \mathcal{K}_{dl}} P_j^{dl} \le P_{tot}^{dl}, \tag{15}$$

$$0 \le P_i^{ul} \le P_{max}^{ul}, i \in \mathcal{K}_{ul}, P_j^{dl} \ge 0, j \in \mathcal{K}_{dl}. \tag{16}$$

As previously discussed, the approach of centralized control can be used to solve the problem at the infrastructure, where the full-duplex BS aggregates comprehensive knowledge across the cellular network. To optimally manage the inter-node interference imposed by the uplink users to downlink receivers, the infrastructure needs the knowledge of all the inter-node interference channel gains. From the form of optimization problem, the required amount of information at the BS is listed in Table I. [2]

| Information at BS | Amount |
|---|---|
| Uplink path loss: $g_i^{ul}$ | All $i \in \mathcal{K}_{ul}$ |
| Downlink path loss: $g_j^{dl}$ | All $j \in \mathcal{K}_{dl}$ |
| Inter-node interference: $g_{ij}^I$ | All $(i,j)$ pairs, $i \in \mathcal{K}_{ul}, j \in \mathcal{N}_i$ |

TABLE I
REQUIRED INFORMATION FOR CENTRALIZED ALGORITHM AT BS

As discussed in the introduction, the above centralized solution is non-scalable primarily because of the significant overhead to potentially collect $K_{ul} \times K_{dl}$ inter-node interference channels. Therefore, our goal is to design a scalable distributed structure, where the network-wide information are not aggregated at the BS. Instead, the information is distributed across the cellular network so that each device only acquires information at most one hop away.

---

[2]Note that in the above table, we do not include the parameters $M$, $P_{tot}^{dl}, P_{max}^{ul}$, which are assumed to be predetermined information that does not need to be communicated network wide.

## V. DISTRIBUTED SOLUTION USING ONE-HOP INFORMATION

In this section, we present a scalable network state information architecture, denoted by one-hop information. Such information architecture significantly reduces the overhead associated with centrally aggregating all the information at the infrastructure. Specifically, the one-hop information architecture is described below in Table II-III, where the Interference-plus-Noise (IN) metric, i.e., $m_{i,j}[t]$ is formally defined in (29) in Section V-C.

| Information at BS | Amount |
|---|---|
| Downlink SINR $\text{SINR}_j^{dl}[t]$ | All $j \in \mathcal{K}_{dl}$ |

TABLE II
ONE-HOP INFORMATION ARCHITECTURE, BS

| Information at uplink user $i$ | Amount |
|---|---|
| Uplink SINR: $\text{SINR}_i^{ul}[t]$ | 1 |
| IN metric: $m_{ij}[t]$ | All $j \in \mathcal{N}_i$ |

TABLE III
ONE-HOP INFORMATION ARCHITECTURE, UPLINK USER $i$

It can be observed from the above tables that the BS requires only SINR from the downlink users, which does not incur additional overhead to current cellular systems where BS periodically collects SINR from downlink users for power and rate control. The BS hence does not need to know information two hops away, e.g., inter-node interference. For each uplink user $i$, it only requires the one-hop information, i.e., the IN metric $m_{ij}$ for all users $j \in \mathcal{N}_i$. It does not require information more than one-hop away, e.g., $m_{i_0 j}[t]$ for all users $i_0 \neq i, j \in \mathcal{K}^{dl}$. Therefore, the multiplicative overhead associated centralized solutions disappears. In the the following sections, we show that under massive MIMO systems with high power gain and hence high SINR, the above one-hop information architecture enables design of optimal distributed power allocation scheme to manage inter-node interference.

### A. High SINR Model

In massive MIMO full-duplex system, as illustrated in the rate expressions (7) and (13), large power gain is brought to each user owing to the large number of antennas at the BS. The high power gain in massive MIMO brings high SINR values at each receivers. Under the high SINR, the uplink transmission rates hence can be approximated by

$$\widetilde{R}_i^{ul}(P_i^{ul}) = \log(\text{SINR}_i^{ul}) = \log\left(\frac{M \cdot P_i^{ul} g_i^{ul}}{N_0}\right), \ i \in \mathcal{K}_{ul}, \tag{17}$$

for $P_i^{ul} \geq P_i^0$ where the parameter $P_0^i$ is the power allocation constraint such that the high SINR approximation is valid, i.e., $\text{SINR}_i^{ul} \gg 1$. Under massive MIMO setup, $P_i^0, i \in \mathcal{K}_{ul}$ are small values. As $M$ scales further up, the value of $P_i^0$ scales down.

In the downlink, the SINR expression given in (12) is high thanks to the large value of $M$ compared to the number of

users in the neighborhood in the range of the cell. Correspondingly, the high SINR model gives the transmission rate expression

$$\widetilde{R}_j^{dl}(P_j^{dl}, \mathbf{P}^{ul}) = \log\left(\frac{M \cdot P_j^{dl} g_j^{ul}}{\sum_{i \in \mathcal{N}_j} g_{ij}^I P_i^{ul} + N_0}\right), \ j \in \mathcal{K}_{dl}, \tag{18}$$

which, similar to the uplink, is valid for $P_j^{dl} \geq P_j^0$. The parameter $P_j^0$ is a small value as the number of BS antennas scales up.

Following the techniques in [6], [30], we perform a change of variables by letting $\widehat{P}_i^{ul} = \log(P_i^{ul})$ and $\widehat{P}_j^{dl} = \log(P_j^{dl})$ for all $i \in \mathcal{K}_{ul}$ and $j \in \mathcal{K}_{dl}$. With the new set of variables and under high SINR approximation (17)(18), the utility maximization problem (14)-(16) becomes

$$\max_{\mathbf{r}} \ \sum_{i \in \mathcal{K}_{ul}} U_i(r_i) + \sum_{j \in \mathcal{K}_{dl}} U_j(r_j) \tag{19}$$

$$s.t., \ \ r_i - \log(Mg_i^{ul}) - \widehat{P}_i^{ul} \leq -\log N_0, \ i \in \mathcal{K}_{ul} \tag{20}$$

$$r_j - \log(Mg_j^{dl}) - \widehat{P}_j^{dl} + \log\left(\sum_{i \in \mathcal{N}_j} g_{ij}^I \exp(\widehat{P}_i^{ul}) + N_0\right) \leq 0,$$

$$j \in \mathcal{K}_{dl} \tag{21}$$

$$\sum_{j \in \mathcal{K}_{dl}} \exp(\widehat{P}_j^{dl}) \leq P_{tot}^{dl} \tag{22}$$

$$\log P_i^0 \leq \widehat{P}_i^{ul} \leq \log P_{max}^{ul}, i \in \mathcal{K}_{ul} \tag{23}$$

$$\widehat{P}_j^{dl} \geq \log P_j^0, j \in \mathcal{K}_{dl}, r_k \geq 0, k \in \mathcal{K}_{ul} \cup \mathcal{K}_{dl} \tag{24}$$

where $\mathbf{r} = [r_k]_{k \in \mathcal{K}_{ul} \cup \mathcal{K}_{dl}}$.

**Observatoin 1:** The high SINR provided by massive MIMO system enables us to transform the non-convex problem (14)-(16) to the above convex optimization problem.

### B. Leveraging Full-Duplex Architecture to Reduce Overhead

One of the main characteristics and bottleneck of the massive MIMO full-duplex system is the inter-node interference. Such interference not only can incur significant overhead in the centralized solution, but also *couples* up- and downlink power control decisions. Because the uplink transmission power causes interference to downlink reception, the downlink power allocation will also be affected accordingly. In this section, we present an interesting observation that there exists a way to *decouple* the interplay between up- and downlink power allocations, thanks to the uni-directional uplink to downlink interference in full-duplex system. This structure helps reduce the overhead associated with obtaining network information to achieve the optimal utility for the aforementioned problem (19)-(24).

For each uplink user $i \in \mathcal{K}_{ul}$, we associate a Lagrangian multiplier $q_i^{ul}$ that corresponds to the constraint (20). A Lagrangian multiplier $q_j^{dl}$ is associated with downlink user $j \in \mathcal{K}_{dl}$ that corresponds to constraint (21). The following

dual function is then obtained

$$D(\mathbf{q}^{ul}, \mathbf{q}^{dl}) = \max_{\mathbf{r}} B(\mathbf{r}, \mathbf{q}^{ul}, \mathbf{q}^{dl}) + \max_{\widehat{\mathbf{P}}^{ul}} V_{ul}(\widehat{\mathbf{P}}^{ul}, \mathbf{q}^{ul}, \mathbf{q}^{dl})$$
$$+ \max_{\widehat{\mathbf{P}}^{dl}} V_{dl}(\widehat{\mathbf{P}}^{dl}, \mathbf{q}^{dl}) - \sum_{i \in \mathcal{K}_{ul}} q_i^{ul} \log N_0. \quad (25)$$

where

$$B(\mathbf{r}, \mathbf{q}^{ul}, \mathbf{q}^{dl}) = \sum_{i \in \mathcal{K}_{ul}} \left( U_i(r_i) - q_i^{ul} \cdot r_i \right) + \sum_{j \in \mathcal{K}_{dl}} \left( U_j(r_j) - q_j^{dl} \cdot r_j \right),$$
$$\tag{26}$$

$$V_{ul}(\widehat{\mathbf{P}}^{ul}, \mathbf{q}^{ul}, \mathbf{q}^{dl}) = \sum_{i \in \mathcal{K}_{ul}} q_i^{ul} \cdot \log(M g_i^{ul}) + \sum_{i \in \mathcal{K}_{ul}} q_i^{ul} \widehat{P}_i^{ul}$$
$$- \sum_{j \in \mathcal{K}_{dl}} q_j^{dl} \cdot \log \Big( \sum_{i \in \mathcal{N}_j} g_{ij}^I \exp(\widehat{P}_i^{ul}) + N_0 \Big),$$
$$V_{dl}(\widehat{\mathbf{P}}^{dl}, \mathbf{q}^{dl}) = \sum_{j \in \mathcal{K}_{dl}} q_j^{dl} \cdot \log(M \cdot g_i^{dl}) + \sum_{j \in \mathcal{K}_{dl}} q_j^{dl} \widehat{P}_j^{dl},$$

with vectors $\mathbf{q}^{ul} = [q_i^{ul}]_{i \in \mathcal{K}_{ul}}$, $\mathbf{q}^{dl} = [q_j^{dl}]_{j \in \mathcal{K}_{dl}}$, $\widehat{\mathbf{P}}^{ul} = [\widehat{P}_i^{ul}]_{i \in \mathcal{K}_{ul}}$, $\widehat{\mathbf{P}}^{dl} = [\widehat{P}_j^{dl}]_{j \in \mathcal{K}_{dl}}$.

Hence, given $\mathbf{q}^{ul}, \mathbf{q}^{dl}$, the dual problem is decomposed into three sub-problems, which leads to the following immediate observation.

**Observation 2.** The uplink power control sub-problem $\max_{\widehat{\mathbf{P}}^{ul}} V_{ul}(\widehat{\mathbf{P}}^{ul}, \mathbf{q}^{ul}, \mathbf{q}^{dl})$, and the downlink power control sub-problem $\max_{\widehat{\mathbf{P}}^{dl}} V_{dl}(\widehat{\mathbf{P}}^{dl}, \mathbf{q}^{dl})$ are *completely decoupled* as separate problems.

**Remark:** The Lagrangian approach uses a similar technique compared with other works that considers power control problems in general networks (e.g., [6]). However, the full-duplex capability at the BS provides a unique decoupling structures for up- and downlink power control. In the full-duplex systems that we have considered, inter-node interference limits system performance, and up- and downlink power controls are coupled because of inter-node interference. The decoupling structure, however, interestingly reveals that these two power control schemes can actually be performed separately, given some pricing information. Such separation structure enables us to significantly reduce the information passing overhead to design the optimal power control policies.

For the uplink and downlink sub-problems, we obtain the gradient expressions as,

$$\frac{\partial V_{ul}(\widehat{\mathbf{P}}^{ul}, \mathbf{q}^{ul}, \mathbf{q}^{dl})}{\partial \widehat{P}_i^{ul}} = q_i^{ul} - \sum_{j: i \in \mathcal{N}_j} q_j^{dl} \frac{g_{ij}^I \exp(\widehat{P}_i^{ul})}{\mathsf{IN}_j}, \quad (27)$$

$$\frac{\partial V_{dl}(\widehat{\mathbf{P}}^{dl}, \mathbf{q}^{dl})}{\partial \widehat{P}_j^{dl}} = q_j^{dl}. \quad (28)$$

The sub-problem $\max_{\mathbf{r}} B(\mathbf{r}, \mathbf{q}^{ul}, \mathbf{q}^{dl})$ is the primal rate adaptation, which can be solved with $r_i^* = (U')^{-1}(q_i^{ul})$, $r_j^* = (U')^{-1}(q_j^{dl})$ where $i \in \mathcal{K}_{ul}, j \in \mathcal{K}_{dl}$. Hence the Lagrangian multipliers can be considered as 'price' for achieving a unit of up- and downlink transmission rate.

We next design a distributed power control algorithm based on the aforementioned intuitions.

## C. Power Control using one-hop Information

We formally present the distributed power control and interference management algorithm next, where the notation $[x]_\mathcal{S}$ stands for projecting the value $x$ onto set $\mathcal{S}$, and the IN metric is

$$m_{ij}[t] = \frac{q_j^{dl}[t] g_{ij}^I}{\mathsf{IN}_j[t]}. \quad (29)$$

---

**Algorithm** Distributed Algorithm for Massive MIMO Full-duplex Systems based on One-Hop Information

---

1: **Initialization phase.** At $t = 0$, set $q_i^{ul}[0] = 0$, $q_j^{dl}[0] = 0$, $\widehat{P}_i^{ul}[0] = \log P_{max}^{ul}$, $\widehat{P}_j^{dl}[0] = \log(P_{tot}^{dl}/K_{dl})$, $\mathsf{IN}_j[0] = N_0$ for all $i \in \mathcal{K}_{ul}$, $j \in \mathcal{K}_{dl}$. Go to step 5.

2: **Downlink power update at the BS.** The BS updates the transmission power to the downlink users according to the gradient (28),

$$\widehat{P}_j^{dl}[t]$$
$$= \left[ \widehat{P}_j^{dl}[t-1] + \gamma \cdot q_j^{dl}[t-1] \right]_{\mathbf{1} \cdot P_j^{dl}[t-1] \leq P_{tot}^{dl}, P_j^{dl}[t-1] \geq P_j^0},$$

where $\gamma$ is the step size.

3: **Uplink power update at the users.** The uplink users update the transmission power to the BS according to the gradient (27),

$$\widehat{P}_i^{ul}[t] = \left[ \widehat{P}_i^{ul}[t-1] + \gamma \cdot \left[ q_i^{ul}[t-1] \right. \right.$$
$$\left. \left. - \sum_{j \in \mathcal{N}_i} m_{ij}[t-1] \exp \left( \widehat{P}_i^{ul}[t-1] \right) \right] \right]_{\log P_0^i \leq \widehat{P}_i^{ul}[t-1] \leq \log P_i^{max}}.$$
$$\tag{30}$$

4: **Price update.** The prices $q_j^{dl}[t], j \in \mathcal{K}_{dl}$ are updated as follows,

$$q_j^{dl}[t] = \left[ q_j^{dl}[t-1] + \gamma \left[ r_j[t-1] - \log \left( \mathsf{SINR}_j[t-1] \right) \right] \right]_{q_j^{dl}[t] \geq 0},$$
$$\tag{31}$$

where $r_j[t] = (U_j')^{-1}(q_j^{dl}[t])$. The uplink user $i \in \mathcal{K}_{ul}$ update its price by

$$q_i^{ul}[t] = \left[ q_i^{ul}[t-1] + \gamma \left[ r_i[t-1] - \log \left( \mathsf{SINR}_i[t-1] \right) \right] \right]_{q_i^{ul}[t] \geq 0},$$
$$\tag{32}$$

where $r_i[t] = (U_i')^{-1}(q_i^{ul}[t])$.

5: $t \rightarrow t + 1$. Go to step 2.

---

In the algorithm, the up- and downlink power is updated according to the gradient (27)(28) in item 2 and 3. The price update in item 4 is along the direction of gradient of dual function $D(\mathbf{q}^{ul}, \mathbf{q}^{dl})$. Also note that in uplink control algorithm (30) in item 3, the term

$$m_{ij}[t-1] \exp \left( \widehat{P}_i^{ul}[t-1] \right) = q_j^{dl}[t-1] \frac{g_{ij}^I P_i^{ul}[t-1]}{\mathsf{IN}_j[t-1]}$$

denotes the contribution of inter-node interference from uplink user $i$ to the overall interference experienced at downlink user $j$, weighted by the price of user $j$. Therefore if user $i$ contributes significant interference to downlink user $j$ whose

achievable rate is already very 'pricey', user $i$'s power will tend to decrease.

Note that the above power control algorithm is implemented distributively at the uplink and downlink users. The distributed nature of the algorithm is facilitated by the decoupling property between up- and downlinks previously discussed, which is captured by the one-hop information. Specifically, because of the decoupling, given the price $\mathbf{q}^{dl}$, downlink power control is implemented independently from uplink users. Hence BS does not require uplink power information. Uplink user power control is implemented autonomously without awareness of the downlink power control policy. Note that the the one-hop information takes advantage of such structure so that all the network devices has sufficient information to perform optimal power control.

The following theorem establishes the convergence of the algorithm. The proof follows the similar lines as [6] and is hence neglected.

**Theorem 1.** *Let $\phi^*$ be the set of $(\widehat{\mathbf{P}}^{ul}, \widehat{\mathbf{P}}^{dl})$ that optimally solves the problem (19)-(24). For small enough $\gamma$, we have $\lim_{t \to \infty} (\widehat{\mathbf{P}}^{ul}[t], \widehat{\mathbf{P}}^{dl}[t]) \in \phi^*$.*

It also follows from [6] that the algorithm converges to optimum geometrically, given in the following theorem, where $(\mathbf{P}^{ul,*}, \mathbf{P}^{dl,*}) \in \phi^*$ and

$$\epsilon[k] = \Big| \sum_{i \in \mathcal{K}_{ul}} U_i(\widetilde{R}_i^{ul}(P_i^{ul}[k])) + \sum_{j \in \mathcal{K}_{dl}} U_i(\widetilde{R}_j^{dl}(P_j^{dl}[k], \mathbf{P}^{ul}[k]))$$
$$- \sum_{i \in \mathcal{K}_{ul}} U_i(\widetilde{R}_i^{ul}(P_i^{ul,*})) + \sum_{j \in \mathcal{K}_{dl}} U_i(\widetilde{R}_j^{dl}(P_j^{dl,*}, \mathbf{P}^{ul,*})) \Big|.$$

**Theorem 2.** *There exist constants $c$ and $\beta$ such that for all $k$, $\epsilon[k] \le c\beta^k$.*

## VI. POWER CONTROL IMPLEMENTATION USING OVERHEARING-BASED SCHEME

We have previously seen how the one-hop information architecture facilitates design of distributed algorithm for managing inter-node interference. Next, we present how the architecture can be implemented using an overhearing-based scheme.

At the end of each slot $t$, a feedback message is transmitted from each downlink user $j \in \mathcal{K}_{dl}$ to the BS. The feedback message $\mathsf{fb}_j[t]$ from the $j$-th downlink user contains the following value,

$$\mathsf{fb}_j[t] = \frac{q_j^{dl}[t]}{\mathsf{IN}_j[t]} = \frac{q_j^{dl}[t] \cdot \mathsf{SINR}_j^{dl}[t]}{MP_j^{dl}[t]g_j^{dl}}, \tag{33}$$

where $\mathsf{IN}_j[t]$ is the interference-plus-noise measured at the $j$-th downlink user. Note also that downlink user $j$ has all the knowledge to keep a copy of price $q_j^{dl}[t]$ according to (32).

The feedback message $\mathsf{fb}_j[t]$ is received at the BS and is overheard by all the uplink users in the neighborhood $N_j$ of $j$-th downlink user. We assume these feedback packets are sent with low rate and can be received collision-free at BS and users via orthogonal frequency/time resources. The BS and the uplink users then process the packets as follows.

• From the embedded pilot symbols in the feedback message $\mathsf{fb}_j[t]$, the BS estimates the downlink path loss $g_j^{dl}$. The BS then uses $\mathsf{fb}_j[t]$ to calculate $\mathsf{SINR}_j^{dl}[t]$ at the downlink receiver $j$ at time $t$, i.e.,

$$\mathsf{SINR}_j^{dl}[t] = \frac{\mathsf{fb}_j[t]MP_j^{dl}[t]g_j^{dl}}{q_j^{dl}[t]}. \tag{34}$$

• The uplink users in the neighborhood $\mathcal{N}_j$ of downlink user $j$ *overhear* the feedback message. From the embedded pilot symbols in the feedback message $\mathsf{fb}_j[t]$, the uplink transmitter $i \in \mathcal{N}_j$ estimates the interference channel gain $g_{ij}^I$ associated with downlink receiver $j \in \mathcal{K}_{dl}$. The uplink transmitters $i \in \mathcal{N}_j$ then, from (29) and (33), use the decoded feedback $\mathsf{fb}_j[t]$ to calculate the metric $m_{ij}[t] = \mathsf{fb}_j[t]P^{ul}[t]g_{ij}^I$ and update the power allocation according to (30).

Therefore, with the overhearing-based feedback scheme, the BS obtains the required SINR information in Table 1, and each uplink user obtains the IN metric in Table 2. Note that in Table 2, uplink SINR can be obtained from (6) where the uplink power is known to the user and path loss coefficient can be estimated via BS broadcasting pilots at initialization to all users.

## VII. ASYMPTOTIC PERFORMANCE OF POWER CONTROL

Massive MIMO full-duplex system brings a large amount of possible degree of freedoms, hence facilitates the capability to simultaneously serve a large number of uplink and downlink users. In this section, we study the performance gain associated with power control, as well as the asymptotic scaling of the optimal power allocation in massive MIMO full-duplex system when both the number of antennas and the number of users scale up.

We hence consider a sequence of scenarios with expanding number of users and antennas, indexed by $l$ with $l \in \mathbb{Z}^+$. For the $k$-th scenario, the set of uplink users, downlink users, and the number of BS antennas are denoted as $\mathcal{K}_{ul}^k$, $\mathcal{K}_{dl}^k$, $M_k$, respectively.

To study the system behavior when more and more users are added to the system, we study the monotonic scenario where $M_l < M_{l+1}$, and the uplink and downlink set of users monotonically expand across scenarios, i.e., $\mathcal{K}_{ul}^1 \subsetneq \mathcal{K}_{ul}^2 \subsetneq \mathcal{K}_{ul}^3 \subsetneq \cdots$ and $\mathcal{K}_{dl}^1 \subsetneq \mathcal{K}_{dl}^2 \subsetneq \mathcal{K}_{dl}^3 \subsetneq \cdots$. We hence have $|\mathcal{K}_{ul}^l| \to \infty$, $|\mathcal{K}_{dl}^l| \to \infty$ and $M_l \to \infty$ as $l \to \infty$.

We assumes the inter-node interference $g_{i.j}^I, i \in \mathcal{K}_{ul}, j \in \mathcal{K}_{dl}$ follow certain distribution with expectation $\mathbb{E}[g_{i,j}^I] = \mathcal{E}$ uniformly across $(i, j)$ pairs. This assumption is valid for the typical scenarios where users randomly appear in geographical locations of a cell, or where inter-node interference arrives via random paths.

We assume the following conditions hold true for the utility functions of the users in the network:

(A1) For each user $k$, $\lim_{r \to \infty} U_k'(r) = 0$, $k \in \mathcal{K}_{ul} \cup \mathcal{K}_{dl}$.

(A2) The utility function $U_i$ of $i$-th user satisfies $U_i \in \mathcal{U}$, where $\mathcal{U} = \{U^1(r), U^2(r), \cdots, U^M(r)\}$ is a finite collection of utility functions.

Assumption (A1) is motivated by the diminishing return idea of the utility functions, and holds true for the well-known $\alpha$-fairness, e.g,

$$U_i(r) = \omega_i \frac{r^{1-\alpha}}{1-\alpha}, \alpha > 0, \alpha \neq 1, \omega_i > 0,$$

which includes many well-known fairness criteria such as proportional fairness, minimum potential delay fairness and max-min fairness as special cases that corresponds to different values of $\alpha$. Assumption (A2) states that there are finite choices of utility functions for each user. These choices correspond to, for instance, different types of data traffic, quality of service requirement, etc.

For ease of exposition, we henceforth use the following notations:

• Vector $\mathbf{P}^{ul,l,*}$ and $\mathbf{P}^{dl,l,*}$ respectively denote the optimal uplink and downlink power allocation that maximizes the network sum-utility (i.e., (19)-(24)) in the $l$-th scenario.

• The utility $U_i^{l,*}$, $i \in \mathcal{K}_{ul}^l \cup \mathcal{K}_{dl}^l$ respectively represents the up- and downlink utility under optimal power allocation $\mathbf{P}^{ul,l,*}$ and $\mathbf{P}^{dl,l,*}$ in the $l$-th scenario.

• The utility $U_i^l(P_{max}^{ul}), i \in \mathcal{K}_{ul}^l$, and $U_j^l(\mathbf{P}_{max}^{ul}), j \in \mathcal{K}_{dl}^l$ represent the utility when each uplink user transmits at maximum power $P_{\max}^{ul}$, and when the downlink users transmits at optimal power that maximizes the sum downlink utility under $\mathbf{P}_{max}^{ul}$.

The next theorem states that when the number $M_l$ of the BS antennas scales at the same order of the multiplication of $|\mathcal{K}_{ul}^l|$ and $|\mathcal{K}_{dl}^l|$, and when the number of downlink users scales at the same order or faster than the number of uplink users, then asymptotically the optimal uplink transmission power scales down.

**Theorem 3.** *Suppose* $\lim_{l \to \infty} \frac{M_l}{|\mathcal{K}_{ul}^l||\mathcal{K}_{dl}^l|} = C$ *for some constant* $C >> 1$, *and* $\limsup_{l \to \infty} \frac{|\mathcal{K}_{ul}^l|}{|\mathcal{K}_{dl}^l|} < \infty$. *For any* $0 < \rho < 1$, *we let* $\Theta_\rho^l \subseteq \mathcal{K}_{ul}^l$ *be the subset of uplink users with* $P_i^{ul,l,*} \leq \rho P_{max}^{ul}$. *Then* $\lim_{l \to \infty} \frac{|\Theta_\rho^l|}{|\mathcal{K}_{ul}^l|} = 1$.

*Proof:* See Appendix A. ∎

**Remark:**

(1) Theorem 3 studies the asymptotic regime when $\lim_{l \to \infty} \frac{M_l}{|\mathcal{K}_{ul}^l||\mathcal{K}_{dl}^l|} = C$, where the BS antennas $M_l$ scales at the same order of the multiplication of $K_{ul}^l$ and $K_{dl}^l$. Intuitively this means the amount of antennas grows sufficiently fast to combat *simultaneously* the growing inter-node interference (from the growing number of uplink users) and also to provide sufficient degrees of freedom to the downlink users to guarantee a non-trivial SINR gain at the downlink users expressed in (12). Theorem 3 also assumes that $\limsup_{l \to \infty} \frac{|\mathcal{K}_{ul}^l|}{|\mathcal{K}_{dl}^l|} < \infty$, i.e., in the asymptotic regime, the number of downlink users scales at the same order or faster than the number of uplink users is valid in typical scenarios where the network is often dominated by the downlink traffic.

(2) Theorem 3 states that in the considered asymptotic regime, the optimal transmission power asymptotically scales down. This is intuitive because the uplink SINR grows linearly with the number of antennas. However, because of the
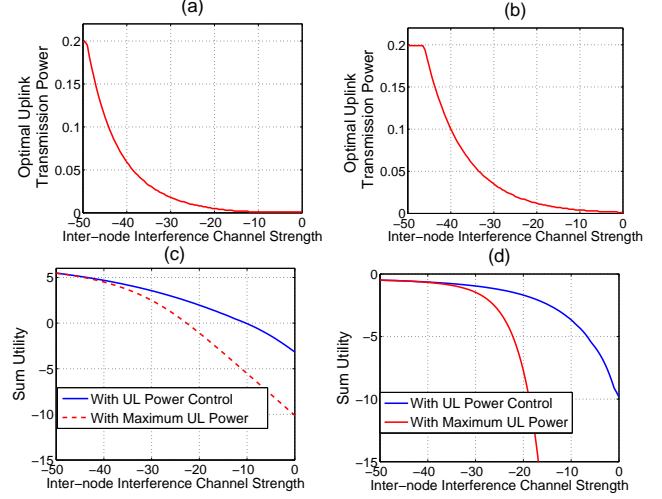


Fig. 2. Performance evaluation of power control algorithm for two-user case. (a) Optimal power scaling with weighted proportional fair utility; (b) Optimal power scaling with weighted minimum potential delay utility;(c) Comparison with Naive scheme with $P^{ul} = P_{max}^{ul}$ under weighted proportional fair utility; (d) Comparison with Naive scheme with $P^{ul} = P_{max}^{ul}$ under weighted minimum potential delay utility.

diminishing decay property of the utility function, the marginal utility gain due to the increasing number of antennas will reduce. Therefore, decreasing the uplink transmission power (e.g., decrease power by $\rho$) will not bring significant loss in the uplink utility, it does, however, bring significant downlink utility gain. This is because, as the number of uplink users increases, the aggregated interference can be large at *each* of the downlink users.

## VIII. NUMERICAL EVALUATIONS

We examine the power control algorithm performance for the massive MIMO full-duplex with one uplink user and one downlink user. We consider the scenario with 128 antennas at the BS, i.e., $M = 128$. The uplink and downlink path loss parameters are $g^{ul} = -60$ $dB$, $g^{dl} = -70$ $dB$ respectively. The power control is subject to the maximum power constraints $P_{max}^{ul} = 23$ $dBm$, $P_{tot}^{dl} = 45$ $dBm$. The noise power is $N_0 = -30$ $dBW$. These parameters are set in accordance with practical values in existing LTE standards. Fig. 2(a) and Fig. 2(c) correspond to weighted proportional fair utility function, with $U^{ul}(r) = \log(r), U^{dl}(r) = 2\log(r)$. Figure 2(b) and Figure 2(d) correspond to weighted minimum potential delay utility function, i.e., $U^{ul}(r) = -1/r, U^{dl}(r) = -2/r$. For this scenario, it is optimal for the downlink user to always transmit at the maximum power since there is only one downlink user.

Figure 2(a)-(b) plot the variation of the optimal uplink transmission power with the gain $g^I$ of the inter-node interference channel. Figure 2(c)-(d) compare the sum-utility under optimal power control and when $P^{ul} = P_{max}^{ul}$. As observed in the figures, when the inter-node interference channel is weak, it is optimal to transmit at the maximum uplink power since the impact on downlink transmission is insignificant. Hence in this regime the uplink transmits power at the maximum power and there is a negligible loss at the downlink as evident in Figure 2(c)-(d). As the inter-node interference channel gain grows stronger, for both cases of utility functions, the
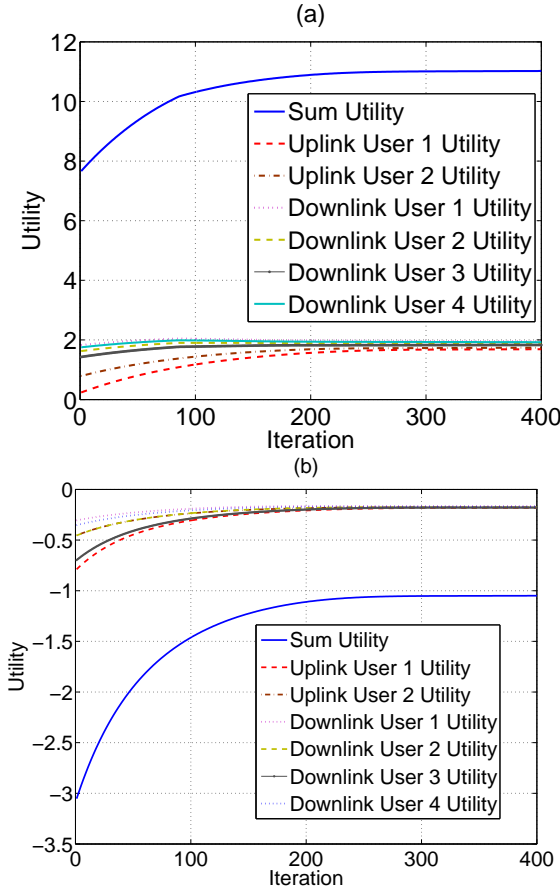
Fig. 3. Convergence of power control algorithm. (a) Proportional fair utility; (b) Minimum potential delay utility.

## IX. Conclusion

In this paper, we study a massive MIMO full-duplex system where the BS contains a large number of full-duplex-capable antennas serving multiple single-antenna half-duplex up and - downlink users. In such system, inter-node inference becomes the main design challenge, where transmission of uplink users creates interference to the reception of downlink users. Because of the vast amount of inter-node interference channels, centralized control by aggregating all network information at the BS will incur significant overhead in the network. We develop an efficient and scalable one-hop information architecture that leverages the unique structure of the massive MIMO full-duplex system. Based on the one-hop information architecture, we propose a distributed power allocation algorithm to optimally manage the inter-node interference, and show how the information can be obtained via a overhearing-based scheme. We characterize the performance of the optimal power control algorithm in the asymptotic regimes where the number of users and the number of antennas at the BS all scale up. We show that power control in large number of BS antennas regime can not only bring utility gains, it also improves the energy efficiency of mobile users.

optimal uplink transmission power decays. This is because under strong interference channel, transmission of uplink users imposes significant interference to the downlink transmission. Because of the diminishing return property of utility functions, reducing the uplink transmission power can alleviate downlink utility loss and bring higher overall sum utility in the network.

Figure 2(c)-(d) also highlight the importance of power control in moderate to high interference regime. As the gain in the inter-node interference channel becomes stronger, power control is increasingly more important to manage the interference created by the uplink users to the downlink receivers to maintain the overall system-level utility, as illustrated by the increasing utility gains in Figure 2(c)(d).

Fig 2(a)-(b) plot convergence performance of the proposed algorithm. We consider the scenario with $128$ antennas at the BS, with $2$ uplink users and $4$ downlink users. The uplink and downlink path loss parameters are $g_1^{ul} = -50dB, g_2^{ul} = -45dB$, and $g_1^{dl} = -56dB, g_2^{dl} = -61dB, g_3^{dl} = -65dB, g_4^{dl} = -58dB$. The inter-node interference channel gain is $g_{1,1}^{I} = -59dB, g_{1,2}^{I} = -60dB, g_{2,1}^{I} = -62dB, g_{2,2}^{I} = -55dB$. Fig 2(a) plots the evolution of the utility function for proportional fairness utilities, i.e., $U_k(r) = \log r, k \in \mathcal{K}_{ul} \cup \mathcal{K}_{dl}$. Fig 2(b) plots the evolution of the utility function for minimum potential delay utility, i.e., $U_k(r) = -1/r, k \in \mathcal{K}_{ul} \cup \mathcal{K}_{dl}$. It can be observed that the algorithm converges at around $200$ iterations.

## References

[1] 3GPP TSG RAN Meeting R1-60.

[2] 3GPP TSG RAN Meeting RP-57.

[3] IEEE 802.11ax-High Efficiency WLAN (HEW) Standardization and Potential Technologies.

[4] D. Bharadia, E. McMilin, and S. Katti. Full duplex radios. In *ACM SIGCOMM Computer Communication Review*, volume 43, pages 375–386. ACM, 2013.

[5] P. Chaporkar and A. Proutiere. Optimal distributed scheduling in wireless networks under sinr interference model. In *Communication, Control, and Computing (Allerton), 2013 51st Annual Allerton Conference on*, pages 1372–1379, Oct 2013.

[6] M. Chiang. Balancing transport and physical layers in wireless multihop networks: Jointly optimal congestion control and power control. *Selected Areas in Communications, IEEE Journal on*, 23(1):104–116, 2005.

[7] M. Chiang, P. Hande, T. Lan, and C. W. Tan. Power control in wireless cellular networks. *Foundations and Trends® in Networking*, 2(4):381–533, 2008.

[8] M. Chiang, C. W. Tan, D. P. Palomar, D. O'Neill, and D. Julian. Power control by geometric programming. *Wireless Communications, IEEE Transactions on*, 6(7):2640–2651, 2007.

[9] N. Choi, M. Patel, and S. Venkatesan. A full duplex multi-channel mac protocol for multi-hop cognitive radio networks. In *Cognitive Radio Oriented Wireless Networks and Communications, 2006. 1st International Conference on*, pages 1–5, June 2006.

[10] B. Di, S. Bayat, L. Song, and Y. Li. Radio resource allocation for full-duplex OFDMA networks using matching theory. In *Computer Communications Workshops (INFOCOM WKSHPS), 2014 IEEE Conference on*, pages 197–198. IEEE, 2014.

[11] M. Duarte, A. Sabharwal, V. Aggarwal, R. Jana, K. Ramakrishnan, C. W. Rice, and N. Shankaranarayanan. Design and characterization of a full-duplex multiantenna system for wifi networks. *Vehicular Technology, IEEE Transactions on*, 63(3):1160–1177, 2014.

[12] E. Everett, A. Sahai, and A. Sabharwal. Passive self-interference suppression for full-duplex infrastructure nodes. *Submitted to Wireless Communications, IEEE Transactions on*, http://arxiv.org/abs/1302.2185, 2013.

[13] C. Guthy, W. Utschick, and M. Honig. Large system analysis of sum capacity in the gaussian mimo broadcast channel. *Selected Areas in Communications, IEEE Journal on*, 31(2):149–159, February 2013.

[14] J. Huang, R. A. Berry, and M. L. Honig. Distributed interference compensation for wireless networks. *Selected Areas in Communications, IEEE Journal on*, 24(5):1074–1084, 2006.

[15] D. Julian, M. Chiang, D. O'Neill, and S. Boyd. Qos and fairness constrained convex optimization of resource allocation for wireless cellular and ad hoc networks. In *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 2, pages 477–486. IEEE, 2002.

[16] Z.-Q. Luo and S. Zhang. Dynamic spectrum management: Complexity and duality. *Selected Topics in Signal Processing, IEEE Journal of*, 2(1):57–73, 2008.

[17] J. Ma and L. Ping. Data-aided channel estimation in large antenna systems. *Signal Processing, IEEE Transactions on*, 62(12):3111–3124, June 2014.

[18] T. Marzetta. Noncooperative cellular wireless with unlimited numbers of base station antennas. *Wireless Communications, IEEE Transactions on*, 9(11):3590–3600, November 2010.

[19] H. Q. Ngo, E. Larsson, and T. Marzetta. Energy and spectral efficiency of very large multiuser mimo systems. *Communications, IEEE Transactions on*, 61(4):1436–1449, April 2013.

[20] L. P. Qian, Y. Zhang, and J. Huang. Mapel: Achieving global optimality for a non-convex wireless power control problem. *Wireless Communications, IEEE Transactions on*, 8(3):1553–1563, March 2009.

[21] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson. Scaling up mimo: Opportunities and challenges with very large arrays. *Signal Processing Magazine, IEEE*, 30(1):40–60, 2013.

[22] A. Sabharwal, P. Schniter, D. Guo, D. Bliss, S. Rangarajan, and R. Wichman. In-band full-duplex wireless: Challenges and opportunities. *Selected Areas in Communications, IEEE Journal on*, 32(9):1637–1652, Sept 2014.

[23] S. Shakkottai, S. G. Shakkottai, and R. Srikant. *Network optimization and control*. Now Publishers Inc, 2008.

[24] S. Shao, D. Liu, K. Deng, Z. Pan, and Y. Tang. Analysis of carrier utilization in full-duplex cellular networks by dividing the co-channel interference region. 2014.

[25] C. Shepard, H. Yu, N. Anand, E. Li, T. Marzetta, R. Yang, and L. Zhong. Argos: Practical many-antenna base stations. In *Proceedings of the 18th annual international conference on Mobile computing and networking*, pages 53–64. ACM, 2012.

[26] C. Shepard, H. Yu, and L. Zhong. Argosv2: A flexible many-antenna research platform. In *Proceedings of the 19th Annual International Conference on Mobile Computing &#38; Networking*, MobiCom '13, pages 163–166, New York, NY, USA, 2013. ACM.

[27] H. Suzuki, I. Collings, D. Hayman, J. Pathikulangara, Z. Chen, and R. Kendall. Large-scale multiple antenna fixed wireless systems for rural areas. In *Personal Indoor and Mobile Radio Communications (PIMRC), 2012 IEEE 23rd International Symposium on*, pages 1600–1605, Sept 2012.

[28] C. W. Tan, M. Chiang, and R. Srikant. Fast algorithms and performance bounds for sum rate maximization in wireless networks. *Networking, IEEE/ACM Transactions on*, 21(3):706–719, 2013.

[29] P. C. Weeraddana, M. Codreanu, M. Latva-Aho, and A. Ephremides. Weighted sum-rate maximization for a set of interfering links via branch and bound. *Signal Processing, IEEE Transactions on*, 59(8):3977–3996, 2011.

[30] Y. Xi and E. M. Yeh. Throughput optimal distributed power control of stochastic wireless networks. *IEEE/ACM Transactions on Networking (TON)*, 18(4):1054–1066, 2010.

[31] H. Yin, D. Gesbert, M. Filippou, and Y. Liu. A coordinated approach to channel estimation in large-scale multiple-antenna systems. *Selected Areas in Communications, IEEE Journal on*, 31(2):264–273, February 2013.

## Appendix A
### Proof of Theorem 2

We let $\Phi_\rho^l = \mathcal{K}_{ul}^l \setminus \Theta_\rho^l$. Hence for all $i \in \Phi_\rho^l$, $P_i^{ul,l,*} > \rho P_{max}^{ul}$. We also let $\tilde{U}'(r) = \max_m\{(U^m)'(r), U^m \in \mathcal{U}\}$.

We prove this lemma by contradiction. Suppose there exist a subsequence $l_m, m = 1, 2, \cdots$, and $0 < \kappa \leq 1$ such that $\lim_{m\to\infty} |\Phi_\rho^{l_m}|/\mathcal{K}_{ul} = \kappa$. We let $\mathbf{P}^{ul,l_m,new}$ be the power allocation vector where $P_i^{ul,l_m,new} = \rho_0 P_{max}^{ul}, i \in \Phi_\rho^{l_m}$ with $\rho_0 < \rho$, and $P_i^{ul,l_m,new} = P_i^{ul,l_m,*}, i \in \Theta_\rho^{l_m}$. For uplink users $i \in \Phi_\rho^{l_m}$ and $m$ being large, the achievable transmission rate

when $P_i^{ul} = P_i^{ul,l_m,new}$ is expressed as follows,

$$R_i^{ul}(P_i^{ul,l_m,new}, M_{l_m}) = \log\left(\frac{M_{l_m} \cdot \rho_0 P_{max}^{ul} g_i^{ul}}{N_0}\right)$$

$$= \log\left(\frac{M_{l_m} \cdot P_i^{ul,l_m,*} g_i^{ul}}{N_0}\right) + \log\left(\frac{M_{l_m} \cdot \rho_0 P_{max}^{ul} g_i^{ul}}{M_{l_m} \cdot P_i^{ul,l_m,*} g_i^{ul}}\right)$$

$$= R_i^{ul}(P_i^{ul,l_m,*}, M_{l_m}) + \log g_i(l_m). \tag{35}$$

where $g_i(l) > \rho_0$. From (35) and concavity of utility function, we have for $i \in \Phi^{l_m}$

$$U_i(R_i^{ul}(P_i^{ul,l_m,new}, M_{l_m})) - U_i(R_i^{ul}(P_i^{ul,l_m,*}, M_{l_m}))$$

$$\geq U_i'(R_i^{ul}(P_i^{ul,new}, M_{l_m})) \cdot \log g_i(l_m). \tag{36}$$

For the downlink users, the achievable transmission rate satisfies

$$R_j^{dl}(P_j^{dl,l_m,*}, \mathbf{P}^{ul,l_m,new}) = \log\left(\frac{M \cdot P_j^{dl,l_m,*} g_j^{dl}}{\sum_{i=1}^{K_{ul}} g_{ij}^I P_i^{ul,new} + N_0}\right)$$

$$= \log\left(\frac{M \cdot P_j^{dl,l_m,*} g_j^{dl}}{\sum_{i \in \Phi^{l_m}} g_{ij}^I \rho_0 P_{max}^{ul} + \sum_{i \in \Theta^{l_m}} P_i^{ul,l_m,*} g_{ij}^I + N_0}\right)$$

$$= R_j^{dl}(P_j^{dl,l_m,*}, \mathbf{P}^{ul,l_m,*}) + \log f_j(l_m), \tag{37}$$

where

$$f_j(l_m) = \frac{\sum_{i \in \mathcal{K}_{ul}^{l_m}} P_i^{ul,l_m,*} g_{ij}^I}{\left(\sum_{i \in \Phi^{l_m}} g_{ij}^I \rho_0 P_{max}^{ul} + \sum_{i \in \Theta^{l_m}} P_i^{ul,l_m,*} g_{ij}^I + N_0\right)}$$

$$= 1 + \frac{\sum_{i \in \Phi^{l_m}} g_{ij}^I P_i^{ul,l_m,*} - \sum_{i \in \Phi^{l_m}} g_{ij}^I \rho_0 P_{max}^{ul} - N_0}{\sum_{i \in \Phi^{l_m}} g_{ij}^I \rho_0 P_{max}^{ul} + \sum_{i \in \Theta^{l_m}} P_i^{ul,l_m,*} g_{ij}^I + N_0},$$

which satisfies,

$$f_j(l_m) \geq 1 + \frac{\left(\sum_{i \in \Phi_\rho^{l_m}} g_{ij}^I [\rho_l - \rho_0] P_{max}^{ul}\right) - N_0}{\left(\sum_{i \in \Phi_\rho^{l_m}} g_{ij}^I \rho P_{max}^{ul} + \sum_{i \in \Theta_\rho^{l_m}} P_i^{ul,l_m,*} + N_0\right)}$$

$$= 1 + \frac{[\rho - \rho_0] P_{max}^{ul} \sum_{i \in \Phi_\rho^{l_m}} g_{ij}^I - N_0}{\left(\sum_{i \in \Phi_\rho^{l_m}} g_{ij}^I \rho P_{max}^{ul} + \sum_{i \in \Theta^l} P_i^{ul,l_m,*} g_{ij}^I\right) + N_0}.$$

Hence

$$\limsup_{m\to\infty} f_j(l_m) \geq 1 + \limsup_{m\to\infty} \frac{[\rho - \rho_0] P_{max}^{ul} \sum_{i \in \Phi^{l_m}} g_{ij}^I / |\mathcal{K}_{ul}^{l_m}|}{\left(\sum_{i \in \mathcal{K}_{ul}^{l_m}} g_{ij}^I \rho P_{max}^{ul}\right) / |\mathcal{K}_{ul}^{l_m}|}$$

$$\geq 1 + \frac{[\rho - \rho_0] C_\kappa}{\rho \mathbb{E}[g_{ij}^I]}.$$

where $C_\kappa = \mathbb{E}[g_{ij}^I | g_{ij}^I \leq \alpha]$ with $\alpha$ being such that $P(g_{ij}^I \leq \alpha) = \kappa$. Next consider the difference in utility. From (37) and concavity of utility, we have for downlink user $j$,

$$U_j(R_j^{dl}(P_j^{dl,l_m,*}, \mathbf{P}^{ul,l_m,new})) - U_j(R_j^{dl}(P_j^{dl,l_m,*}, \mathbf{P}^{ul,l_m,*}))$$

$$\geq U_j'(R_j^{dl,l_m,*}(P_j^{dl,l_m,*}, \mathbf{P}^{ul,l_m,new})) \cdot \log f_j(l_m). \tag{38}$$

We let $\mathbf{P}^{dl,l_m,new*}$ denote the optimal downlink power allocation to maximize the sum-downlink-utility given uplink

power allocation $\mathbf{P}^{ul,l_m,new}$. The difference in overall sum-utility is

$$\left[\sum_{i\in\mathcal{K}_{ul}^{l_m}}U_i^{l_m,new}+\sum_{j\in\mathcal{K}_{dl}^{l_m}}U_j^{l_m,new,*}\right]-\left[\sum_{i\in\mathcal{K}_{ul}^{l_m}}U_i^{l_m,*}+\sum_{j\in\mathcal{K}_{dl}^{l_m}}U_j^{l_m,*}\right]$$

$$\geq\left[\sum_{i\in\mathcal{K}_{ul}^{l_m}}U_i^{l_m,new}+\sum_{j\in\mathcal{K}_{dl}}U_j^{l_m,new}\right]-\left[\sum_{i\in\mathcal{K}_{ul}^{l_m}}U_i^{l_m,*}+\sum_{j\in\mathcal{K}_{dl}^{l_m}}U_j^{l_m,*}\right]$$

$$=\left[\sum_{i\in\mathcal{K}_{ul}^{l_m}}U_i^{l_m,new}-U_i^{l_m,*}\right]+\left[\sum_{j\in\mathcal{K}_{dl}^{l_m}}U_j^{l_m,new}-U_j^{l_m,*}\right]$$

$$\geq\sum_{i\in\Phi^{l_m}}U_i'(R_i^{ul}(P_i^{ul,l_m,new},M_{l_m}))\cdot\log g_i(l_m)$$

$$+\sum_{j\in\mathcal{K}_{dl}^{l_m}}U_j'(R_j^{dl,l_m,*}((P_j^{dl},\mathbf{P}^{ul,l_m,new}))\cdot\log f_j(l_m).$$

where the last inequality holds from (36)(38). We then have

$$\liminf_{m\to\infty}\left[\sum_{i\in\Phi^{l_m}}U_i'(R_i^{ul}(P_i^{ul,l_m,new},M_{l_m}))\cdot\log g(l_m)\right.$$

$$\left.+\sum_{j\in\mathcal{K}_{dl}^{l_m}}U_j'(R_j^{dl}((P_j^{dl,l_m,*},\mathbf{P}^{ul,l_m,new}))\cdot\log f_j(l_m)\right]$$

$$\geq\liminf_{m\to\infty}\left[|\Phi_\rho^{l_m}|\widehat{U}'(R_i^{ul}(P_i^{ul,l_m,new},M_{l_m}))\log\rho_0\right.$$

$$\left.+\sum_{j\in\mathcal{K}_{dl}^{l_m}}U_j'(R_j^{dl}((P_j^{dl,l_m,*},\mathbf{P}^{ul,l_m,new}))\cdot\log f_j(l_m)\right],$$

where recall that $\widehat{U}'(r)=\max_i U_i'(r)$. Note that

$$\limsup_{m\to\infty}R_j^{dl}(P_j^{dl,l_m,*},\mathbf{P}^{ul,l_m,new})$$

$$=\limsup_{m\to\infty}\log\left(\frac{M\cdot P_j^{dl,l_m,*}g_j^{dl}}{\sum_{i\in\Phi^{l_m}}g_{ij}^I\rho_0 P_{max}^{ul}+\sum_{i\in\Theta^{l_m}}P_i^{ul,l_m,*}g_{ij}^I+N_0}\right)$$

$$\leq\limsup_{m\to\infty}\log\left(\frac{M\cdot P_j^{dl,l_m,*}g_j^{dl}}{\sum_{i\in\Phi^{l_m}}g_{ij}^I\rho_0 P_{max}^{ul}+\sum_{i\in\Theta^{l_m}}P_i^{ul,l_m,*}g_{ij}^I+N_0}\right)$$

$$\leq\limsup_{m\to\infty}\log\left(\frac{M\cdot P_j^{dl,l_m,*}g_j^{dl}/|\mathcal{K}_{ul}^{l_m}|}{\left(\sum_{i\in\Phi^{l_m}}g_{ij}^I\rho_0 P_{max}^{ul}+N_0\right)/|\mathcal{K}_{ul}^{l_m}|}\right)$$

$$=\limsup_{m\to\infty}\log\left(\frac{M\cdot P_j^{dl,l_m,*}g_j^{dl}/|\mathcal{K}_{ul}^{l_m}|}{C_\kappa\mathbb{E}[g_{ij}^I]\rho_0 P_{max}^{ul}}\right), \tag{39}$$

where recall that $C$ was defined in the statement of the theorem, and $C_\kappa$ was previously defined in this proof. We proceed with the following lemma.

**Lemma 1.** *Let* $\Psi_\omega^l=\{j\in\mathcal{K}_{dl}^l:P_j^{dl,l,*}<\omega\frac{P_{tot}^{ul}}{|\mathcal{K}_{dl}^l|}\}$ *for* $\omega>0$. *There exists* $\omega>0$ *such that* $\limsup_{l\to\infty}\frac{|\Psi_\omega^l|}{|\mathcal{K}_{dl}^l|}>0$.

*Proof:* Suppose the statement is not true. Then for any $\omega>1$ we have $\lim_{l\to\infty}\frac{|\Psi_\omega^l|}{|\mathcal{K}_{dl}^l|}=0$. Therefore

$$\lim_{l\to\infty}\sum_{j\in\mathcal{K}_{dl}^l}P_j^{dl,l,*}\geq\lim_{l\to\infty}\sum_{j\notin\Psi_\omega^l}P_j^{dl,l,*}\geq\omega P_{tot}^{ul}>P_{tot}^{ul}.$$

Hence the total downlink transmission power exceeds the power constraint, thus establishing a contradiction. ∎

According to the above lemma, we let $\omega_0$ and $\varepsilon$ be such that $\limsup_{m\to\infty}\frac{|\Psi_{\omega_0}^{l_m}|}{|\mathcal{K}_{dl}^{l_m}|}=\varepsilon$. Therefore

$$\liminf_{m\to\infty}\left[|\Phi_\rho^{l_m}|\widehat{U}'(R_i^{ul}(P_i^{ul,l_m,new},M_{l_m}))\log\rho_0\right.$$

$$\left.+\sum_{j\in\mathcal{K}_{dl}^{l_m}}U_j'(R_j^{dl}((P_j^{dl,l_m,*},\mathbf{P}^{ul,l_m,new}))\cdot\log f_j(l_m)\right],$$

$$\geq\liminf_{m\to\infty}\left[|\Phi_\rho^{l_m}|\widehat{U}'(R_i^{ul}(P_i^{ul,l_m,new},M_{l_m}))\log\rho_0\right.$$

$$\left.+|\Psi_{\omega_0}^{l_m}|U_j'\left(\frac{C\omega_0 g_j^{dl}}{C_\kappa\mathbb{E}[g_{ij}^I]\rho_0}\right)\cdot\log f_j(l_m)\right],$$

$$=K_{ul}^{l_m}\liminf_{m\to\infty}\left[\kappa\widehat{U}'(R_i^{ul}(P_i^{ul,l_m,new},M_{l_m}))\log\rho_0\right.$$

$$\left.+\omega_0\frac{|\mathcal{K}_{dl}^{l_m}|}{|\mathcal{K}_{ul}^{l_m}|}U_j'\left(\frac{C\omega_0 g_j^{dl}}{C_\kappa\mathbb{E}[g_{ij}^I]\rho_0}\right)\cdot\log f_j(l_m)\right],$$

$$>0,$$

where the last equality holds because the first term in the limit decays to 0 with $l_m$ of (39), $\widehat{U}'(r),U'(r)$ monotonically decrease with $r$, the path loss $g_j^{dl}<1$, and because from the theorem statement $\liminf_{m\to\infty}\frac{|\mathcal{K}_{dl}^{l_m}|}{|\mathcal{K}_{ul}^{l_m}|}>0$.

Therefore choosing $(\mathbf{P}^{ul,l,new},\mathbf{P}^{dl,l,new*})$ asymptotically achieves utility higher than the optimal power allocation $(\mathbf{P}^{ul,l,*},\mathbf{P}^{dl,l,*})$, which contradicts to $(\mathbf{P}^{ul,l,*},\mathbf{P}^{dl,l,*})$ being the optimal power allocation. We hence proved the Theorem.